Structural bioinformatics

VDNA: The virtual DNA plug-in for VMD

Thomas C. Bishop

Center for Computational Science, Tulane University, Lindy Boggs Center Suite 500, New Orleans, LA 70118, USA

ABSTRACT

Summary: The DNA inter base pair step parameters (Tilt, Roll, Twist, Shift, Slide, Rise) are a standard internal coordinate representation of DNA. In the absence of bend and shear, it is relatively easy to mentally visualize how Twist and Rise generate the familiar double helix. More complex structures do not readily yield to such intuition. For this reason, we developed a plug-in for VMD that accepts a set of mathematical expressions as input and generates a coarse-grained model of DNA as output. This feature of VDNA appears to provide a unique approach to DNA modeling. Predefined expressions include: linear, sheared, bent and circular DNA, and models of the nucleosome superhelix, chromatin, thermal motion and nucleosome unwrapping.

Availability: VDNA is pre-installed in VMD, http://www.ks.uiuc.edu/Research/vmd. Updates are at http://dna.ccs.tulane.edu.

Contact: bishop@tulane.edu

1 INTRODUCTION

The DNA inter base pair step helical parameters (Tilt, Roll, Twist, Shift, Slide, Rise) are widely employed to model the structure, energetics and dynamics of DNA. These parameters describe the relative rotations and translations along a length of DNA needed to move from one base pair to the next. The parameters are well defined (Dickerson, 1989), and there are a number of programs for extracting these parameters from a Cartesian coordinate representation of DNA or for assembling a model of DNA given a set of helical parameters (El Hassan and Calladine, 1995; Lavery and Sklenar, 1988; Lu and Olson, 2003; Macke and Case, 1998; Tung and Carter, 1994; Vlahovicek and Pongor, 2000). However, none of these enable the user to specify a set of mathematical functions for the helical parameters and rapidly visualize the corresponding 3D model. This is a major obstacle in development of an understanding of the complex relations between local and global structure. Only for the most trivial distributions of DNA helical parameters is it easy to mentally visualize the global structure of DNA. For this purpose, we have developed a plug-in for VMD (Humphrey *et al.*, 1996) that allows users to generate models of dsDNA using arbitrarily complex mathematical expressions.

2 METHODS

The inter base pair step parameters are a local or internal coordinate representation of DNA that specify how a set of directors attached to the *i*-th base pair and denoted $(\hat{d}_1, \hat{d}_2, \hat{d}_3)$ are rotated (Tilt, Roll, Twist) and translated (Shift, Slide, Rise) to achieve the set of directors associated with

the $(i+1)$-th base pair. The local bend is ($\sqrt{\text{Roll}^2 + \text{Tilt}^2}$) and local shear is ($\sqrt{\text{Shift}^2 + \text{Slide}^2}$). In Cartesian coordinates, DNA is a space curve with centerline $\vec{r}(s_i)$ and directors $\hat{d}_k(s_i)$ that results from the accumulation of bend, shear, Twist and Rise as a function of s_i. We emphasize that the helical parameter description is suitable for any fiber-like molecule for which a set of local directors can be embedded in the fiber. Thus, while VDNA is intended to represent DNA it is actually much more widely applicable. The complete set of helical parameters includes intra base pair parameters (Shear, Stretch, Stagger, Buckle, Propeller-Twist, Opening), which describe deformations of the base pairs. VDNA does not utilize these descriptors.

Here, let $\vec{\Omega} = (\text{Tilt}, \text{Roll}, \text{Twist})$ and $\vec{\Gamma} = (\text{Shift}, \text{Slide}, \text{Rise})$ represent the DNA helical parameters. In general $\vec{\Omega}$ and $\vec{\Gamma}$ are functions of s and t where s represents the continuum limit of position measured in base pair steps along the DNA and t represents time. To obtain a 3D representation of DNA as a space curve $\vec{r}$ with local directors $\hat{d}_k$ requires integration of $\vec{\Omega}(s,t)$ and $\vec{\Gamma}(s,t)$ as a function of s at time t. VDNA simply provides an interface for defining $\vec{\Omega}(s,t)$ and $\vec{\Gamma}(s,t)$, integrating the expressions and displaying results. In case $\vec{\Omega}$ and $\vec{\Gamma}$ are time dependent, VDNA automatically integrates $\vec{\Omega}$ and $\vec{\Gamma}$ from $s=0$ to $s=\max S$ for each instance $t=0$ to $\max T$ with a time step of 1. The result is $\max T + 1$ structures corresponding to $t=0,1,2,\ldots,\max T$.

For the purposes of numerical integration VDNA utilizes El Hassan's algorithm (El Hassan and Calladine, 1995). This algorithm has proven to be extremely fast and robust and ensures invariance to the direction of integration, i.e. integration from $s=0$ to $\max S$ or $s=\max S$ to 0 gives the same result. Note the sign of Shift and Tilt must be changed during a 'reverse' integration.

3 USAGE

The main VDNA panel, Figure 1, is generated by selecting 'Extensions → Visualization → Virtual DNA Viewer' from the VMD pull-down menus. This panel provides text boxes for inputing mathematical expressions for the helical parameters. These strings are used to create the new procedures `omega` and `gamma` which are defined as functions of the variables *s* and *t*. As the structure at time *t* is created the variable *s* varies from 0 to max *S*. In case max *T* > 0, VDNA will generate a total of max *T* + 1 structures. In all cases, the angular deformations are defined in degrees per base pair step, and the translations are given in ångström per base pair step. Any string that forms a valid argument to the tcl command 'expr' can be utilized in defining the helical parameter statements. (Consult the examples or a tcl language reference source for more details.) VDNA recognizes several defined variables: V1, V2, Nuc and Lk ($Pi is predefined as π to single precision). The procedures `omega` and/or `gamma` utilize whatever text is entered for these variables. Such variable substitution provides a convenient means of introducing arbitrarily complex expressions into more than one helical parameter statement simultaneously. The text box 'Cores' defines the procedure
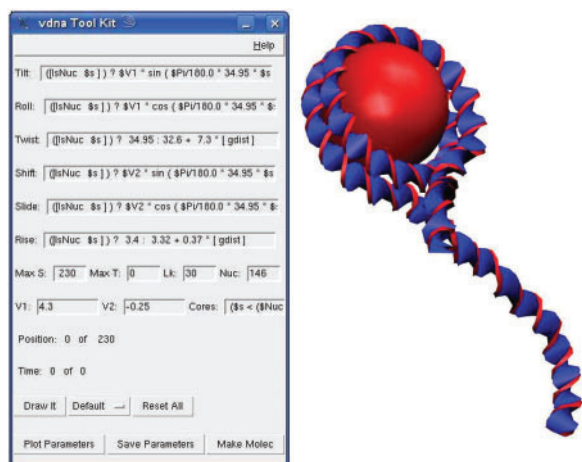
IsNuc which is called during the main integration loop with the argument s. When IsNuc $= 1$, VDNA maintains a running average of $\vec{r}(s)$. When IsNuc switches to zero, VDNA draws a sphere at the accumulated position and resets the average. This function is useful for indicating proteins docked to the DNA, e.g. the nucleosome's histone core.

Finally, VDNA includes a number of predefined examples from which to choose, see Figure 2. Some examples are intended to demonstrate geometric relations and to assist in developing intuition rather than model biologic entities. These models are not pictured, but include: straight, bent, sheared, polygonal, circular and superhelical structures. Others are designed to be building blocks for modeling biologically relevant structures, Figure 2, but are not parameterized to represent specific biophysical entities. The default model, Figure 1, demonstrates how to combine the generic, predefined Shear Helix and Thermal models to accurately represent a specific biological entity: a Shear Helix model of nucleosome core particle (Bishop, 2008) with DNA linkers that exhibit average B-DNA conformation and fluctuation properties (Lavery *et al.*, 2009).

4 FUTURE PLANS AND CONCLUSION

Features to be included in future versions of VDNA include sequence-dependent values for the helical parameters, greater support for conversion between all atom models and helical parameter models using either 3DNA or Curves and the reading of helical parameters from file. Presently, the parameter file produced by VDNA can be used to create an all atom model with 3DNA.
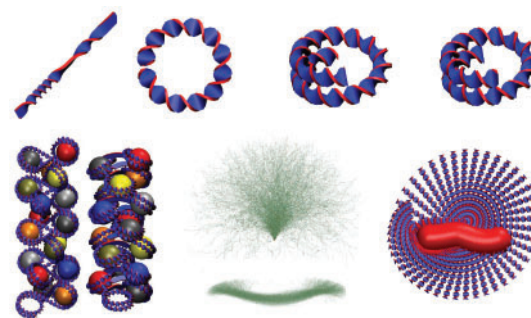
We have introduced a general purpose tool for creating fiber-like models, four-atom-per-base-pair models and all atom models of DNA from mathematical expressions. Predefined expressions are provided as building blocks to investigate structural, thermal and dynamical properties including: shear (Shift or Slide); non-uniform Twist; a Torsion Helix and a Shear Helix (Bishop, 2008); linker DNA in simple models of chromatin; thermal variations; and time evolutions. VDNA is written entirely in tcl so it can be readily modified by users with limited programming experience.

REFERENCES

Bishop,T.C. (2008) Geometry of the nucleosomal DNA superhelix. *Biophys J.*, **95**, 1007–1017.

Dickerson,R.E. (1989) Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.*, **17**, 1797–1803.

El Hassan,M.A. and Calladine,C.R. (1995) The assessment of the geometry of dinucleotide steps in double-helical DNA; a new local calculation scheme. *J. Mol. Biol.*, **251**, 648–664.

Humphrey,W. *et al.* (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38, 27–28.

Lavery,R. and Sklenar,H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.

Lavery,R. *et al.* (2009) A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.*, [Epub ahead of print, doi:10.1093/nar/gkp834, October 22, 2009].

Lu,X.-J. and Olson,W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.

Macke,T. and Case,D.A (1998) Modeling unusual nucleic acid structures. In Leontes,N.B. and SantaLucia,J. Jr., eds. *Molecular Modeling of Nucleic Acids*, American Chemical Society, Washington, DC, pp. 379–393.

Tung,C S. and Carter,E.S. (1994) Nucleic acid modeling tool (NAMOT): an interactive graphic tool for modeling nucleic acid structures. *Comput. Appl. Biosci.*, **10**, 427–433.

Vlahovicek,K. and Pongor,S. (2000) Model.it: building three dimensional DNA models from sequence data. *Bioinformatics*, **16**, 1044–1045.